Individual Assignment 5: Design

Chris Anderson
PREDICT 455 Section 56

Summary and Problem Definition

Included in the "Originals" is a series of graphics I developed for a personal research project to assess the performance of MLB pitchers using data from the PITCHf/x motion capture system. The graphics were presented individually along with paragraphs of text explaining their meaning and implications. The goal of this exercise is to improve and condense these graphics into a data visualization dashboard that tells a story on its own without additional explanation.

Methods

PITCHf/x is a motion tracking system that captures data on the release point, speed, and trajectory of every pitch that occurs in every game throughout the league. Some time ago, the statistical software R was used to analyze the data and produce the original graphics. The data was manipulated into two datasets, which were extracted from R into .csv format and loaded into Tableau to produce the dashboard.

Programming Overview

In Tableau, three individual worksheets were created and combined into a dynamic dashboard. General design best practices were followed throughout the development. Titles, labels, gridlines, and annotations were colored in an understated grey shade so as not to take away from the main point of the graphics. Minimal accents were used throughout the dashboard, and color was only used to distinguish groups in a categorical variable. The same color scheme was used consistently across multiple plots to avoid confusion. Size was used as a meaningful dimension, but elements such as angles or volume were avoided because they are proven to be difficult to perceive accurately. Width, length,

and area were used to denote certain measurements instead. Brief annotations throughout the dashboard explain key graphical elements.

Results

The result is a clean, visually appealing dashboard that incorporates several new data elements and tells a more complete story than the original graphics. The line graphs show pitch velocity and spin rate throughout the pitcher's outing to illustrate the effects of fatigue, if any. Each pitch type is represented by a different colored line, so users can see how well the pitcher mixed in all of the various pitches available to him. The width of the line corresponds to the ball-strike count at the time the pitch was thrown. Coaches using the dashboard can look at this element in search of any troubling pitch selection patterns that may benefit the opposing batters.

The scatterplot shows the exit velocity and launch angle for every batted ball allowed by the pitcher(s), including a "Barrel Zone" to highlight the most impactful combinations of exit velocity and launch angle. The color of each data point corresponds to the pitch type in the same manner as the line graphs; the shape of the points indicate whether the ball was a hit or an out; and the size of the point represents its hang time. These dimensions allow users to asses each pitcher's performance on a deeper level than simply looking at hits and runs allowed. For example, Pitcher X allowed three hits in the Barrel Zone, but two of them were outs. This suggests that he was the beneficiary of some luck because these types of batted balls frequently result in home runs and run-producing extra-base hits. Over time, his results may suffer if he continues to allow such hits. Finally, the bar chart displays various metrics on a per-plate-appearance scale to show a summary of the pitcher's efficiency and overall performance.

The dashboard is filterable by pitcher so the user can review the team as a whole or focus on each pitcher individually. One PDF file is included for each pitcher to demonstrate the filtering functionality. The entire Tableau workbook is also included for users to interact with the dashboard.